REDUCING THE EFFECTS OF AN UNCONTROLLED VARIABLE IN THE OPTIMIZATION OF A QUADRATIC RESPONSE SURFACE

James C. Craig, Jr.
T. Frederick Holden
Nicholas C. Aceto
and
Edwin F. Schoppet

2988

In the event that control and optimization of a process require a mathematical model, and no adequate theoretical model exists, the best alternative may be an empirical quadratic response surface. The justification for this generally is that the relation between an input variable and the output is nonlinear, and furthermore that the input variables interact in their action on the output. A good approach is the use of Box-Wilson rotatable composite experimental designs (*1*) and subsequent estimation of the coefficients describing the response surface by least squares. When a description of the process is arrived at in these terms, graphical methods may be used to locate a desirable region of operation, up to four or five independent variables. Where more variables must be considered, a number of computer search techniques exist for locating this region. Hoerl of DuPont has developed a method (*2*, *3*) of relating both dependent and independent variables at an extremum to the radius of the hypersphere centered on the experimental design center. The result of this analysis is the set of coordinates for the maximum and minimum ridges in terms of the process inputs and the corresponding levels of output variable, all given as a function of this radius. It has the advantages of showing both absolute and local maxima and minima in a two-dimensional form and constraining the solutions to reasonable levels.

A notable complication is the requirement that a desirable extremum be located in a region where a second dependent variable is at or exceeds a specified level, for example, maximum product rate, given that purity is higher than 99%. Hoerl has extended his ridge analysis to include the constrained solution. In this case, the maximum ridges are used at a specified radius as the initial point in a search for a maximum for one dependent variable, given a fixed value for a second dependent variable.

A variant of the problem is the case where an independent variable can be measured but not controlled. Here again, a response surface study is possible. However, the preceding methods discussed for optimization cannot be applied directly. It is of no value to have a specified input for a variable which cannot be controlled. An obvious approach to this is to find, if they exist, coincident areas of the two response surfaces under investigation where changes due to variation in the uncontrolled variable are minimal, and a constrained optimum within that region. That is, each response is limited to the region where the effect of the uncontrolled variable is lowest, and the hypersurface defined by the intersection of the two limited response surfaces is examined. This paper discusses the mathematical and experimental methodology employed in an example of this problem.

Eastern Utilization Research and Development Division. T. Frederick Holden is with The Lubrizol Corporation, Wickliffe, Ohio.

EXPERIMENTAL PROCEDURE

During the development of a vacuum foam-drying process for whole milk (*4* to *6*), it was noted that an apparently uncontrollable seasonal variation in the milk supply had a marked effect on product rate and quality. Foam subsidence in the dryer and the consequent residual foam thickness largely determines drying rate. A laboratory test (*7*) was developed to measure foam stability. A parameter calculated from this test was found to characterize dryer performance. It varied with season, demonstrating greater foam stability in summer and less in winter. It also varied significantly from day to day within seasons. It was decided that process conditions must be sought which would yield a relatively constant output at an economically feasible rate during all seasons.

Prior to the study discussed here, nondesigned experimentation disclosed eleven controllable variables in the dryer which had a significant effect on product moisture and quality. Many or all of these were expected to interact with the seasonal variable. Consequently, all of these were studied when experimental designs were initiated.

The experimental year can be divided roughly into four seasons. During summer and winter foam stability shows no significant trend with time. During spring and fall, the two transition seasons, a gross change was found over a 2- to 3-mo. period. An additional limitation in the experimentation was that it was planned to use data, as acquired, to evolve toward the general region of higher product rates and lower moisture content. This required that interim results be of a form which would allow interpretation.

In order to determine the effects of the controlled variables and their interactions, it was necessary to limit an experiment to one season. As no more than one and one-half experimental points could be taken per day (two-day processing cycle), only small fractional factorial experiments with composite points (*8*, *9*) could be run. During constant seasons $\frac{1}{2}\,2^6$ or $\frac{1}{2}\,2^7$ was used as the base design. It was found early in the experimentation that an appreciable number of quadratic or interaction terms were significant. Consequently, all further work was planned to include estimates of the second-degree terms. The uncontrolled variable (θ) could not be studied during the constant season.

The strategy used to study θ was as follows. Transition (spring and fall) experiments were blocked whole or fractional 3^n factorial designs, with replicated zero points. For example, four blocks, each of a $\frac{1}{3}(3^4)$ factorial design, were run with the center point repeated at intervals. The first and last blocks were the same. The intent was that independent estimates of the four controllable variables studied could be made, while day to day variation in θ would have a low correlation with these estimates. The gross change in θ through blocks allows an estimate of the effects and interactions of θ.

MATHEMATICAL TREATMENT

In all cases, a second-degree equation was used, and an estimate of the coefficients was formed by multiple regression. Only the significance of the entire model was tested. While factorial or composite designs were used throughout, no attempt was made to use coefficients derived from an analysis which assumed orthogonality. Measurable deviations from the design occurred frequently. The designs on which the experiments were based, however, led to an essentially diagonal matrix. The dependent variables were moisture content and 5-hydroxymethylfurfural (HMF) content [a measure of heat damage (*6*, *10*)]. Starting with these models, the problem was to minimize moisture for a specified low level of HMF and a given product rate with minimal variation due to θ.

Each equation has the form

$$Y = Y_0 + \sum_{i=1}^{11} \alpha_i X_i + \sum_{i=1}^{11} \sum_{j=1}^{11} \alpha_{ij} X_i X_j + \theta \sum_{i=1}^{11} \alpha_{i\theta} X_i + \alpha_\theta \theta + \alpha_{\theta\theta}\,\theta^2 \qquad (1)$$

In order to achieve a solution, the coefficients of θ must be made negligible. Prior to dropping θ from the equations, its effect was minimized. This was done by selecting two values of θ which represent practical extremes. An intermediate value of θ is then found which would maximize the variation in Y due to the linear and quadratic coefficients of θ, evaluated at these three points. A new value for the linear coefficient of θ is found to best fit these points, giving a linear approximation over this range. The θ^2 term is now dropped. When the Y equation is differentiated with respect to θ, the new linear coefficient becomes a constant. The resulting equation is solved for an independent variable and substituted into the original equation which now describes a response surface, reduced by two dimensions, on which the effect of the uncontrolled variable is minimal. In this case, substitution was made in both the Y_1 and Y_2 equations, with $\partial Y_1/\partial\theta = A_1$, and then another variable was eliminated in both by $\partial Y_2/\partial\theta = A_2$. The resultant equations, now constrained to minimize variation in both Y_1 and Y_2 due to θ, were treated alternately as constrained and constraining equations by using Hoerl's ridge analysis (*2*, *3*).

Two of the independent variables, belt loading (pounds per square feet) and residence time (minutes), determine product rate. In order that a series of solutions have the same product rate, a rate and a series of belt loading values were preselected. These choices determined residence times, so that substitution was made for two more variables before the analysis.

RESULTS AND DISCUSSION

Figure 1 shows a plot of percent moisture vs. HMF at two radii for the HMF minimization, constrained by moisture. Hoerl's original solution, and his computer program, which was used, maximized the constrained variable. Consequently, the equation to be maximized was the negative of HMF. There is a desirable region where HMF is less than or equal to the maximum acceptable, and a point is chosen in that region, with the choice based on the

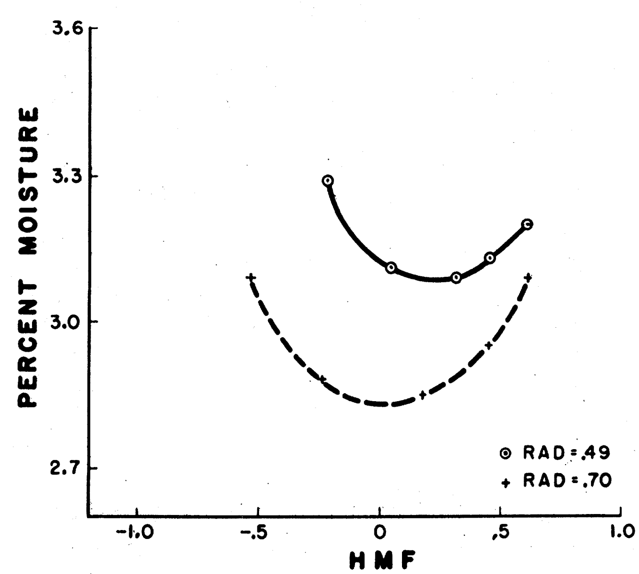

Fig. 1. Moisture content vs. HMF content for constrained solution at 2 radii.

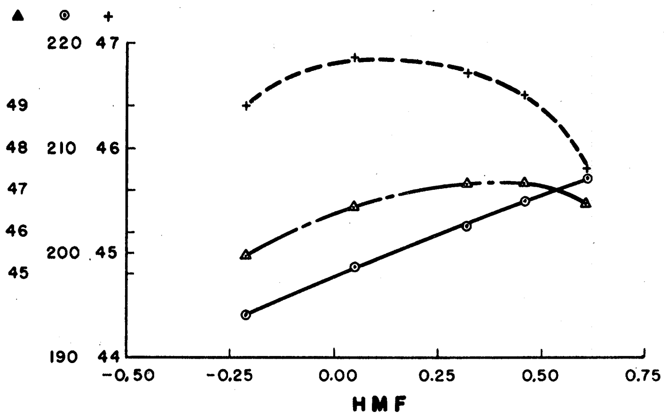

Fig. 2. Independent variables for ridge analysis solution vs. HMF.

relative importance of moisture level and ease of operation. Figure 2 shows some of the independent variables also plotted against HMF. From these results, a set of conditions was chosen and tested in the process. Over the limited range of seasonal variation thus far tested, good agreement between the predicted and observed results was found.

While it can be concluded that this approach was successful, a number of points should be noted.

1. The attempt to minimize and then neglect the effect of the uncontrolled variable will not necessarily work. One can predict, after the solution, the residual variation, but if it is too large, nothing more can be done.

2. The final solution for a maximum contained an implicit six constraints. At some radii no ridge could be found, and at others the only results had no physical meaning. It seems quite possible that in some situations so many constraints would permit only imaginary solutions.

3. The attempt to constrain by substitution negates some of the simplicity of Hoerl's method. The two variables substituted to account for seasonal variation must be solved for after the ridge solution. Trial and error is required to find a range of radii corresponding to reasonable values of the implicit variables.

An amount of judgment is required to use the mathematical procedure described. Values of θ, radius, and the substituted variables must be chosen on a basis of technical feasibility and expected behavior of the process.

CONCLUSIONS

The validity of this procedure must be established by predicting operating conditions which can be tested experimentally. The procedure described did, in fact, yield reasonable conditions for a moderately complex process. Owing to the nonexplicit form of analysis required, it is hoped that this work will encourage a more rigorous mathematical development.

ACKNOWLEDGMENT

The authors acknowledge the contributions of A. E. Hoerl and R. J. Hunn of E. I. DuPont de Nemours & Company for the ridge analysis program and their advice on its use and interpretation of the results. They also acknowledge the support and guidance of R. K. Eskew, as well as the contribution of E. S. Della Monica and his staff for the chemical analyses.

LITERATURE CITED

1. Box, G. E. P., and K. B. Wilson, *J. Roy. Stat. Soc.*, **B13**, 1 (1951).
2. Hoerl, A. E., *Chem. Eng. Progr.*, **55**, 69 (1959).
3. Draper, N. R., *Technometrics*, **5**, 469 (1963).
4. Sinnamon, H. I., N. C. Aceto, R. K. Eskew, and E. F. Schoppet, *J. Dairy Sci.*, **40**, 1036 (1957).
5. Aceto, N. C., H. I. Sinnamon, E. F. Schoppet, and R. K. Eskew, *ibid.*, **45**, 501 (1962).
6. Schoppet, E. F., N. C. Aceto, R. K. Eskew, J. C. Craig, Jr., and T. F. Holden, *ibid.*, **48**, 1436 (1965).
7. Holden, T. F., N. C. Aceto, and E. F. Schoppet, *ibid.*, **47**, 359 (1964).
8. Davies, O. L., ed., "The Design and Analysis of Industrial Experiments," 2 ed., Hafner, New York (1956).
9. Cochran, W. G., and G. M. Cox, "Experimental Designs," 2 ed., Wiley, New York (1957).
10. Craig, J. C., Jr., N. C. Aceto, and E. S. Della Monica, *J. Dairy Sci.*, **44**, 1827 (1961).